Use of Semiempirical Methods for Detecting Anomalies in Reported Enthalpies of Formation of Organic Compounds

James J. P. Stewart[a)]

Stewart Computational Chemistry, 15210 Paddington Circle, Colorado Springs, Colorado 80921

(Received 28 January 2003; revised manuscript received 24 July 2003; accepted 25 August 2003; published online 13 August 2004)

Using semiempirical methods, a survey of the enthalpies of formation of ~1300 compounds in the NIST Chemistry WebBook database was performed and compared to the reported values. Five sets of compounds were identified: a large set in which theory and experiment agreed, a set in which they disagreed strongly due to transcription errors, a set in which they disagreed strongly due to experimental errors, a set that disagreed strongly due to errors in the method, and a set in which the difference between calculation and experiment was not obvious. © *2004 American Institute of Physics.*
[DOI: 10.1063/1.1643403]

Key words: enthalpies of formation; errors; semiempirical methods; thermochemistry.

Contents

List of Tables

[a)]Electronic mail: jstewart@us.fujitsu.com

List of Figures

1. Introduction

The degree of accuracy of the prediction of properties calculated using computational chemistry methods can be determined by a comparison of the computed values with those obtained by experiment. Such a comparison requires that a sufficiently large collection of data exists so that the results are of statistical significance, and, more important, that the experimental data are reliable. Determining the accuracy of experimental data, particularly thermochemical data, is normally the province of the experimentalist. However, while performing routine comparisons of calculated and experimental enthalpies of formation, severe differences were observed. These differences were considerably larger than the probable error assigned to the experimental data, and were also inconsistent with the errors expected from the computational method. In an attempt to resolve this anomaly, the probable value of each datum was estimated using experimental quantities to estimate either enthalpies of reaction for hypothetical reactions or by comparison with related compounds. In each case, the experimental value was also inconsistent with that predicted using experimental values. The only conclusion that could be made was that the datum in question was inaccurate.

If, indeed, several reference data in a collection of such data were inaccurate, then the validity of determining the accuracy of a semi-empirical method by comparison with that collection becomes questionable. This poses a severe problem, in that no alternatives exist. To resolve this, a systematic comparison was made of a large number of experimentally determined enthalpies of formation with the results of calculations. In this paper, the results of that comparison are presented.

1.1. NIST Database

The NIST database[1] is a large Web-based (http://webbook.nist.gov/chemistry) collection of experimentally determined enthalpies of formation and other data. Such a collection is obviously of use to both experimentalists and theoreticians interested in comparing the predictions of their models with what is known. In general, each such entry includes the experimental enthalpy of formation, a measure of the accuracy, and the original reference. Several collections of thermochemical data exist, however the NIST database is the most convenient and comprehensive of the readily available compendia. Because of this, it has proven invaluable for both experimentalists and theoreticians. For this work, a subset of the collection was used that consisted only of those organic compounds for which enthalpies of formation in the gas phase at 298.15 K were available.

1.2. Types of Errors

1.2.1. Experimental Errors

An important distinction must be made regarding errors. Of their nature, experimental errors in enthalpies of formation are random. A minor exception might be the case where unintentional bias was introduced by the experimentalist. This could happen if, for example, the experimental value did not agree with previously reported values for related species. If no faults were found in the experimental procedure, then the experimental value should be reported. If the value was not reported, the absence of the value introduces a bias into the collection.

1.2.2. Computational Errors

Of their nature, computational errors are systematic. Errors can be assigned to functional groups and to structural elements, such as rings, methylene units, and branches. For example, in the *n*-alkanes, the experimentally observed change in enthalpy of formation in going to the next higher homolog rapidly converges on -20.5 kJ mol^{-1}. In common semi-empirical methods, this quantity varies from -28.9 kJ mol^{-1} in MNDO[2,3] and AM1[4] to -21.8 kJ mol^{-1} in PM5.[5] Some computational errors can be random; this is a consequence of a mistake being made. For example, the orientation of a hydroxyl group is normally regarded as being unimportant; the assumption being made that during geometry optimization any errors in initial orientation would be automatically corrected. However, this assumption relies on the premise that only one orientation is stable. In several instances, unexpected disagreement was found between the computed and experimental enthalpies of formation. These were traced to a hydroxyl group being orientated in a high-energy conformation. Upon correcting the geometric error, the enthalpy of formation decreased to the expected value. All random errors of this type are necessarily positive.

2. Theory

This work was prompted by problems encountered while developing a semi-empirical method for use in biochemistry. The applicability of such a method could be limited to well-

behaved systems that are stable *in vivo*, that is, in aqueous media. This limitation determined the composition of the reference data set used in this work. The set of data used consisted of over 1300 compounds from the *NIST WebBook*. In order to get a representative sample, the size of system ranged from methane (five atoms) to 5,6-dibutyl-5,6-bis(4-*tert*-butylphenyl)decane (100 atoms). The range of elements allowed was limited to H, C, N, O, F, S, Cl, Br, and I. Only compounds that contained carbon and one or more of these elements were considered in this work. Compounds for which the reported error was large, e.g., greater than about 6 kJ mol^{-1}, were excluded, as were compounds with very large positive enthalpies of formation. This latter set was excluded because of their assumed high reactivity and presumed instability in aqueous media. In rare instances, some compounds with very large positive enthalpies of formation, such as 1,3,5 tricyanobenzene, were included. All such molecules were presumed to be stable because of specific functional groups.

2.1. Computational Method Used

There are several commonly available semi-empirical methods; these include MNDO,[2,3] AM1,[4] PM3,[6,7] and PM5.[5] These methods share a common algebraic form and differ mainly in the numerical values of the parameters. For this work, none of these methods was considered suitable, because the average error in enthalpies of formation for well-behaved systems of the type considered here was relatively large. This was a consequence of the fact that these methods were designed to handle a wide range of systems, from molecules to ions to transition states. Instead, the parameter set was reoptimized for the specific set of molecules of interest here. In the optimization, the weighting factors for enthalpies of formation and geometry were increased relative to the weighting factors for other properties. Parameter optimization for semi-empirical methods is a purely mathematical procedure, and is described elsewhere.[6,8] All calculations reported here use the standard semi-empirical algebraic form, but with the parameter set defined here. Because the parameter set was designed for a single use—to reproduce enthalpies of formation with increased accuracy—it should not be regarded as a new method suitable for general use.

For convenience, the CAChe ProjectLeader program[9] was used. This allowed large numbers of geometries and enthalpies of formation to be calculated automatically. Analyses of the results were performed using Microsoft EXCEL.

A measure of accuracy is provided by the distribution of differences between the experimental and calculated enthalpies of formation. This distribution is shown in Table 1, and the average error of various semi-empirical methods is shown in Table 2. In determining average errors, reference data identified in this paper as being unreliable were excluded from consideration.

Examination of the errors revealed that some were associated with specific functional groups. Such errors, being systematic, were attributed to faults in the computational model. For the current work, a summary of these is presented in Table 3.

TABLE 1. Distribution of errors

Deviation from experiment (kJ mol^{-1})	Percentage of compounds[a]	Number of compounds
< 5	28.4	357
<10	46.0	577
<15	61.2	768
<20	73.5	923
<25	81.4	1022
<30	86.6	1087
<40	92.3	1158

[a]Set of 1255 compounds.

TABLE 2. Average errors in enthalpies of formation for various semi-empirical methods (all values in kJ mol^{-1})

Method	Median unsigned error	Average unsigned error	Root mean square error
MNDO	28.9	61.3	124.6
AM1	27.4	40.6	58.2
PM3	19.2	24.9	33.3
PM5	15.4	20.9	28.9
This work	11.1	16.1	24.4

3. Comparison of Calculated and Experimental ΔH_f

The distribution of errors is of interest in that the largest contribution to the average error arises from comparatively few compounds. As expected, most errors are relatively small, as shown in Fig. 1 and Table 1. It is reasonable to assume that in those instances where calculation and experiment agree, both the method and the reference data are accurate; this assumption will be used extensively here in Sec. 3. Conversely, where they disagree, either the experimental ΔH_f or the results of the calculation are faulty. From this it follows that any large errors in experimental data are likely to be found where the difference between calculation and observation is largest.

3.1. Errors in the Theoretical Method

Three functional groups were found to be poorly predicted by the semi-empirical method: they are listed in Table 3. No clear reason for the errors in the azo or the alkynes was

TABLE 3. Errors in method

Functional group	Error (kJ mol^{-1})
Alkyne	−41.8
Disulfone	+418.4
Azo	+62.8

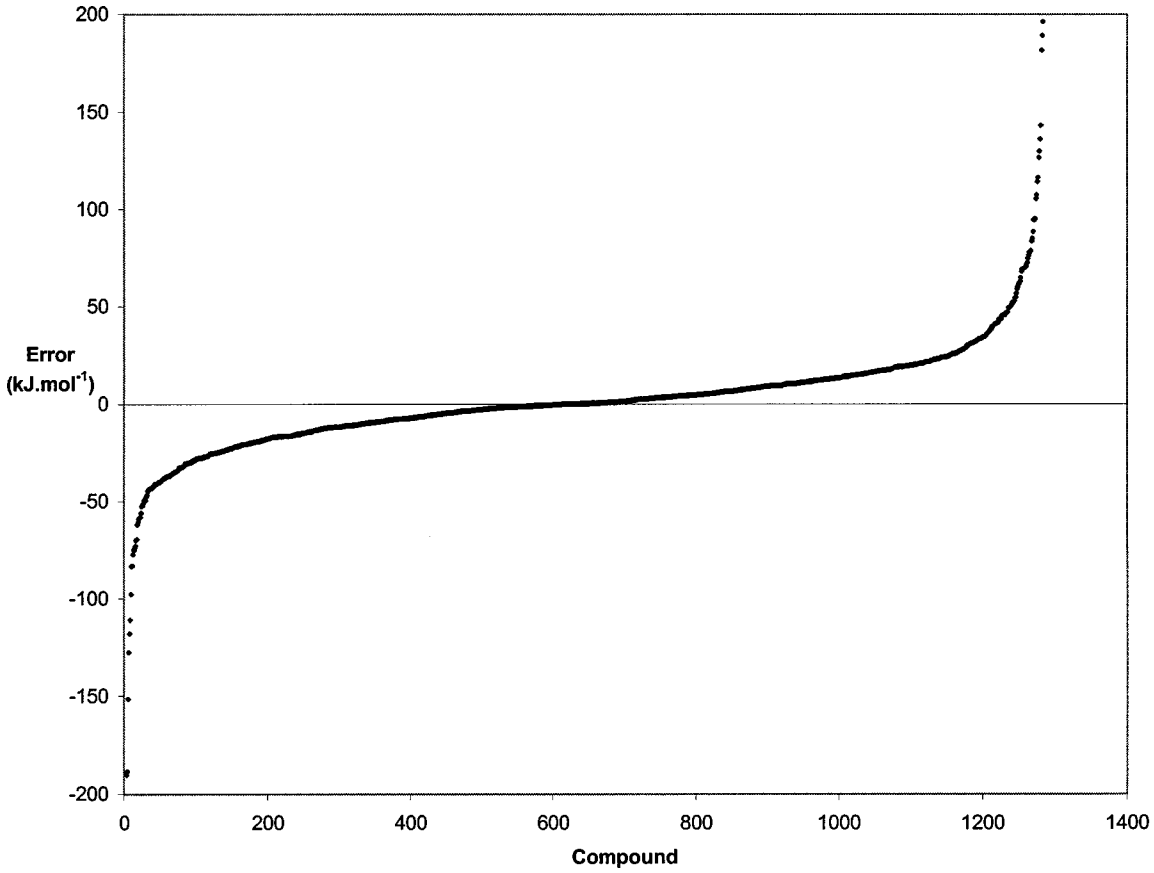

FIG. 1. Distribution of errors.

.

identified; however, in the case of diphenyl disulfone (the only disulfone used), the large error was probably due to the neglect of *d* orbitals.

3.2. Errors in Reference Data

Possible errors in reference data can be grouped into three sets, each of which will be addressed in the following discussion. Evidence of experimental error was derived using experimental data for related compounds, rather than relying on the results of semi-empirical calculations. In all cases, evidence of the accuracy of those data was provided by the agreement with semi-empirical results.

TABLE 4. ΔH_f of polyfluorinated *n*-butane derivatives (all values in kJ mol^{-1})

Compound	Formula	Expt ΔH_f	Calc ΔH_f	Difference
1,3-Perfluorobutadiene	C_4F_6	−942.2	−971.5	−29.3
2-Perfluorobutene	C_4F_8	−946.0	−1565.7	−619.7
Perfluorobutane	C_4F_{10}	−2233.8	−2133.0	100.8

3.3. Transcription Errors

Errors in transcription should not be regarded as ''experimental error,'' but the effect of such errors is the same as if they were experimental errors. Only two such errors were identified.

3.3.1. 2-Perfluorobutene

Although 2-perfluorobutene appears to be an unremarkable compound, the published enthalpy of formation is inconsistent with the related compounds, 1,3-perfluorobutadiene and perfluorobutane, as shown in Table 4. Addition of a fluorine molecule to an unsaturated hydrocarbon is usually associated with a large release of energy. However, when perfluorobutene is formed by addition of F_2

TABLE 5. ΔH_f for 2-*tert*-butyl-*para*-cresol and related compounds (all values in kJ mol^{-1})

Compound	Expt ΔH_f	Calc ΔH_f	Difference
2-*Tert*-butyl-*para*-cresol	+207.0	−209.2	−416.2
Para-cresol	−125.3	−130.5	−5.2
Tert-butyl benzene	−22.7	−13.4	+9.3
Benzene	+82.9	+87.4	+4.5

to 1,3-perfluorobutadiene, only 3.8 kJ mol^{-1} of energy is released. Conversely, when F_2 is added to 2-perfluorobutene to give perfluorobutane, 1287.8 kJ mol^{-1} of energy is released. These quantities are unrealistic. Semi-empirical calculations indicate that the enthalpy of formation of 2-perfluorobutene should be about halfway between that of the other two compounds.

3.3.2. 2-*Tert*-Butyl-*Para*-Cresol

Enthalpies of formation of 2-*tert*-butyl-para-cresol and related compounds are shown in Table 5, from which the following hypothetical reaction can be constructed:

tert-butylbenzene + *para*-cresol

⇒ benzene + 2-*tert*-butyl-*para*-cresol.

The observed and calculated enthalpies of reaction are 437.9 and 22.1 kJ mol^{-1}, respectively.

A cursory comparison of the reported and calculated values indicates a severe anomaly in the enthalpy of formation of 2-*tert*-butyl-para-cresol, and the obvious conclusion is that the reported value is in error by a simple sign.

3.4. Possible Errors in Experiment

3.4.1. Octafluorotoluene

Experimental and calculated ΔH_f for octafluorotoluene, $C_6F_5-CF_3$, and related compounds are shown in Table 6. Using these data, the enthalpy of reaction for the hypothetical reaction,

2,3,4,5,6-pentafluorotoluene + trifluoromethylbenzene

⇒ octafluorotoluene + toluene

can be estimated; the experimental and computed ΔH_r are +204.9 and +33.0 kJ mol^{-1}, respectively. An estimate of the enthalpy of reaction of a half-fluorinated compound disproportionating into the fully fluorinated compound and the corresponding hydrocarbon can be made using the related reaction,

2 1,1,1-trifluoroethane ⇒ ethane + hexafluoroethane.

For this reaction, the experimental and computed ΔH_r are +69.7 and +76.2 kJ mol^{-1}, respectively. Based on these results, the experimental enthalpy of formation of octafluorotoluene is unexpectedly high.

TABLE 6. ΔH_f for octafluorotoluene and related compounds (all values in kJ mol^{-1})

Compound	Formula	Expt ΔH_f	Calc ΔH_f	Difference
Octafluorotoluene	C_7F_8	−1269.4	−1492.9	−223.5
Hexafluoroethane	C_2F_6	−1343.9	−1334.3	+9.6
2,3,4,5,6-Pentafluorotoluene	$C_6F_5-CH_3$	−843.3	−881.6	−38.3
Trifluoromethylbenzene	$C_6H_5-CF_3$	−581.0	−593.3	−12.3
Toluene	C_7H_8	+50.0	+51.0	+1.0
Ethane	C_2H_6	−83.8	−75.7	+8.1
1,1,1-Trifluoroethane	CH_3-CF_3	−748.7	−743.1	+5.6

TABLE 7. ΔH_f for undecylcyclohexane and related compounds (all values in kJ mol^{-1})

Compound	Formula	Expt ΔH_f	Calc ΔH_f	Difference
Cyclopentane	C_5H_{10}	−76.4	−89.5	−13.1
Cyclohexane	C_6H_{12}	−123.1	−120.9	+2.2
Decane	$C_{10}H_{22}$	−249.7	−246.9	+2.8
Undecane	$C_{11}H_{24}$	−270.3	−268.2	+2.1
Undecylcyclopentane	$C_5H_9-C_{11}H_{23}$	−312.5	−329.3	−16.8
Decylcyclohexane	$C_6H_{11}-C_{10}H_{21}$	−336.9	−337.2	−0.3
Undecylcyclohexane	$C_6H_{11}-C_{11}H_{23}$	−148.1	−358.6	−210.5

TABLE 8. ΔH_f for 1,3,5-tricyanobenzene and related compounds (all values in kJ mol^{-1})

Compound	Formula	Expt ΔH_f	Calc ΔH_f	Difference
Benzene	C_6H_6	+82.9	+87.4	+4.5
Benzonitrile	$C_6H_5(CN)$	+219.0	+220.5	+1.5
Ortho-dicyanobenzene	$C_6H_4(CN)_2$	+367.5	+367.8	+0.3
Meta-dicyanobenzene	$C_6H_4(CN)_2$	+362.7	+361.5	−1.3
Para-dicyanobenzene	$C_6H_4(CN)_2$	+358.3	+362.3	+4.0
1,3,5-Tricyanobenzene	$C_6H_3(CN)_3$	+313.0	+509.6	+196.6

TABLE 9. ΔH_f for diethyl malonate and related compounds (all values in kJ mol^{-1})

Compound	Expt ΔH_f	Calc ΔH_f	Difference
Dimethyl malonate	−737.8	−743.5	−5.7
Diethyl malonate	−921.7	−787.4	+134.3
Methyl acetate	−410.0	−405.0	+5.0
Ethyl acetate	−445.4	−427.6	+17.8
Methane	−74.9	−59.0	+15.9

3.4.2. Undecylcyclohexane

That the experimental enthalpy of formation of undecylcyclohexane is inconsistent with that of closely related compounds can readily be shown by comparison with such compounds, as shown in Table 7.

3.4.3. 1,3,5-Tricyanobenzene

Addition of a cyano group to benzene is normally associated with an increase in energy of about 134 kJ mol^{-1}. In going from the dicyanobenzenes to the symmetric tricyanobenzene, the experimental enthalpy of formation decreases by 50 kJ mol^{-1} (Table 8) but the calculations suggest that the trend observed should continue, that is, the enthalpy of formation should increase by about 150 kJ mol^{-1}.

3.4.4. Diethyl Malonate

The calculated and experimental enthalpies of formation of diethyl malonate differ by 134 kJ mol^{-1} but the enthalpy of formation of the closely related dimethyl malonate was accurately predicted (Table 9). In order to determine whether the change in enthalpy of formation should be attributed to the change in going from the methyl to the ethyl ester, esters of acetic acid were examined. Using the experimental change in enthalpies of formation in going from methyl to ethyl ester, the expected change in going from dimethyl malonate to diethyl malonate was 71 kJ mol^{-1}. The experimental change was 184 kJ mol^{-1}.

This is unexpectedly large, so if it follows that either the enthalpy of formation of the dimethyl or diethyl ester should be regarded as suspect. The observed and calculated enthalpies of reaction for

$$2 \text{ ethyl acetate} \Rightarrow \text{diethyl malonate} + \text{methane}$$

are −106 and +8.8 kJ mol^{-1}, respectively. This reaction is unlikely to involve a large enthalpy of reaction. Both this result and the results of semi-empirical calculations suggest that the dimethyl ester value was correct, and that the diethyl value was incorrect.

3.4.5. Bis-(*n*-Perfluoropropyl) Ether

The evidence of a problem in the experimental enthalpy of formation of bis-(*n*-perfluoropropyl) ether is indirect. From the experimental enthalpies of formation of perfluoroheptane and hexafluoroethane, Table 10, an estimate can be made of the contribution of a CF_2 group, −408.3 kJ mol^{-1}, from which an estimate can be made of the enthalpy of formation of *n*-perfluorohexane, −2977 kJ mol^{-1}. Using the experimental enthalpies of formation of hexafluorodimethyl ether and hexafluoroethane, an estimate can be made of the energy released when an oxygen atom is inserted into a perfluorinated C–C bond: 200 kJ mol^{-1}, from which the enthalpy of formation of bis-(*n*-perfluoropropyl) ether can be approximated as −3177 kJ mol^{-1}.

Many highly fluorinated compounds have larger than average errors in the predicted enthalpy of formation. Also, the assumption has been made that the energy of insertion of an oxygen atom into a perfluorinated C–C bond is not significantly different between C_2F_6 and C_6F_{14}. Because of this, the analysis given here is less rigorous and less accurate than desirable.

3.4.6. 2,6,6-Trimethyl-2-Cyclohexen-1-One

An estimate of the enthalpy of formation of 2,6,6-trimethyl-2-hexen-1-one can be made by relating it to simpler compounds. The structure can be regarded as a polymethylated cyclohexenone; in turn, cyclohexenone can be regarded as cyclohexane with two functional groups: an unsaturated bond and a ketone.

In going from cyclohexane to cyclohexene, the change in energy observed is +119.1 kJ mol^{-1} (Table 11); this was used as an estimate of the energy contribution of the olefin group. In going from cyclohexane to cyclohexone, the experimental change is −107.7 kJ mol^{-1}. In turn, this is an estimate of the energy contribution of the ketone group. Using these quantities, the enthalpy of formation of cyclohexenone can then be estimated as −112.0 kJ mol^{-1}. The addition of methyl groups is associated with a drop in energy of at least 29 kJ mol^{-1}, therefore the enthalpy of formation of 2,6,6-trimethyl-2-cyclohexen-1-one can be estimated to be less than −199 kJ mol^{-1}. Based on this estimate, the experimental value of −44.2 kJ mol^{-1} is unreasonably large.

TABLE 10. ΔH_f for bis-(*n*-perfluoropropyl) ether and related compounds (all values in kJ mol^{-1})

Compound	Formula	Expt ΔH_f	Calc ΔH_f	Difference
Bis-(*n*-perfluoropropyl) ether	$C_6F_{14}O$	−3105.0	−3218.8	−113.8
n-Perfluoroheptane	C_7F_{16}	−3385.4	−3391.1	−5.7
Hexafluorodimethyl ether	C_2F_6O	−1543.9	−1576.1	−32.2
Hexafluoroethane	C_2F_6	−1343.9	−1334.3	+9.6

TABLE 11. ΔH_f for 2,6,6-trimethyl-2-cyclohexen-1-one and related compounds (all values in kJ mol^{-1})

Compound	Expt ΔH_f	Calc ΔH_f	Difference
2,6,6-Trimethyl-2-cyclohexen-1-one	−44.2	−232.6	−188.4
Cyclohexone	−231.1	−251.0	−19.9
Cyclohexene	−4.3	−20.9	−16.6
Cyclohexane	−123.4	−120.5	+2.9
1-Methyl cyclohexene	−81.3	−58.6	+22.7
Methyl cyclohexane	−154.8	−147.7	+7.1

TABLE 12. ΔH_f for 1,2,3,6-tetrahydropyridine and related compounds (all values in kJ mol^{-1})

Compound	Expt ΔH_f	Calc ΔH_f	Difference
Benzene	+82.9	+87.4	+4.5
Cyclohexa-1,3-diene	+104.6	+75.7	−28.9
Cyclohexene	−4.3	−20.9	−16.6
Cyclohexane	−123.4	−120.5	+2.9
Pyridine	+140.2	+124.7	−15.5
1,2,3,6-Tetrahydropyridine	+140.6	+29.7	−110.9
Piperidine	−47.2	−67.8	−20.6

TABLE 13. ΔH_f for isophthalamide and related compounds (all values in kJ mol^{-1})

Compound	Formula	Expt ΔH_f	Calc ΔH_f	Difference
Benzene	C_6H_6	+82.9	+87.4	+4.5
Benzamide	$C_6H_5(CONH_2)$	−100.9	−98.7	+2.2
Isophthalamide	$C_6H_4(CONH_2)_2$	−382.0	−294.1	+87.9
Terephthalamide	$C_6H_4(CONH_2)_2$	−376.0	−291.6	+84.4

TABLE 14. ΔH_f for 4-hydroxy-4-methylpentan-2-one and related compounds (all values in kJ mol^{-1})

Compound	Formula	Expt ΔH_f	Calc ΔH_f	Difference
4-Hydroxy-4-methylpentan-2-one	$C_6H_{12}O_2$	−540.7	−486.2	+54.5
Tert-butanol	$C_4H_{12}O$	−312.5	−308.8	+3.7
Acetone	C_3H_6O	−218.5	−232.6	−14.1
Methane	CH_4	−74.9	−59.0	+15.9

3.4.7. 1,2,3,6-Tetrahydropyridine

When benzene is hydrogenated, the first step, that is, to form cyclohexadiene, involves an increase in energy, which is attributable to the loss of aromaticity. The next two steps, to form first cyclohexene, then cyclohexane, are accompanied by a large drop in energy, as shown in Table 12. The electronic structure of pyridine is similar to that of benzene; therefore, by analogy, the energy changes upon hydrogenation should be similar.

Unfortunately, the experimental enthalpy of formation of the dihydropyridines does not appear to be reported, however the heat of formation of the tetrahydro- and fully hydrogenated pyridines is available. As with benzene, when four hydrogen atoms are added to pyridine, the calculated enthalpy of formation drops significantly (95.0 kJ mol^{-1} for pyridine, 108.3 kJ mol^{-1} for benzene). However, the reported change is a decrease in heat of formation of only 0.4 kJ mol^{-1}. This seems unusually small, and is in contrast with the reported change in enthalpy of formation when the last two hydrogen atoms are added, a drop of 187.9 kJ mol^{-1}.

3.4.8. Isophthalamide and Terephthalamide

The experimental enthalpies of formation of both benzene and benzamide are reproduced with good accuracy (Table 13) therefore the change upon adding the $-CONH_2$ group is likely to be modeled correctly. This change is −183.8 kJ mol^{-1}, experimentally, and −186.1 kJ mol^{-1}, calculated. Experimentally, the thermochemical effect of adding a second $-CONH_2$ group to form isophthalamide and terephthalamide is to lower the enthalpy of formation by 281.1 and 275.1 kJ mol^{-1}, respectively. Computationally, the decreases were 195.4 and 192.9 kJ mol^{-1}. The large decrease observed experimentally is unexpected, and casts doubt on the reported enthalpy of formation of the phthalamides.

3.4.9. 4-Hydroxy-4-Methylpentan-2-One

Although there are no enthalpies of formation reported for compounds that are similar to 4-hydroxy-4-methylpentan-2-

TABLE 15. ΔH_f for 2,2′-diquinoline and related compounds (all values in kJ mol^{-1})

Compound	Formula	Expt ΔH_f	Calc ΔH_f	Difference
Benzene	C_6H_6	+82.9	+87.4	+4.5
Biphenyl	$C_{12}H_{10}$	+182.0	+182.8	+0.8
Hydrogen	H_2	0.0	−9.6	−9.6
Pyridine	C_5H_5N	+140.2	+124.7	−15.5
2,2′-Dipyridine	$C_{10}H_8N_2$	+267.9	+269.4	−1.5
Quinoline	C_9H_7N	+200.5	+200.8	+0.3
2,2′-Diquinoline	$C_{18}H_{12}N_2$	+347.9	+425.9	+78.0

one, the presence of the insulating CH_2 group allows a reaction to be constructed, a reaction for which the expected enthalpy of reaction should be very small:

tert-butanol + acetone

⇒4-hydroxy-4-methylpentan-2-one + methane.

Using experimental enthalpies of formation (Table 14) the enthalpy of reaction is predicted to be −84.6 kJ mol^{-1}. This is unexpectedly large, and in strong disagreement with the calculated enthalpy of reaction of 3.8 kJ mol^{-1}. The experimental and calculated enthalpies of formation for three of the four species agree well, but the experimental and calculated values for the fourth species, 4-hydroxy-4-methylpentan-2-one, differ by 54.5 kJ mol^{-1}.

3.4.10. 2,2′-Diquinoline

The experimental enthalpy of reaction of pyridine ring coupling to form 2,2′-dipyridine is −12.5 kJ mol^{-1} (Table 15) whereas computationally the value is +10.4 kJ mol^{-1}. For quinoline, the experimental enthalpy of reaction to form 2,2′-diquinoline is −53.1 kJ mol^{-1}. The computed value is considerably different, +14.7 kJ mol^{-1}, and is similar to the value for the dimerization of pyridine. The reported value of the enthalpy of formation of 2,2′-diquinoline appears to be too negative by about 40–75 kJ mol^{-1}.

3.4.11. 6-Methyl-5-Hepten-2-One

Although 6-methyl-5-hepten-2-one is a relatively simple ketone, the computed and experimental enthalpies of formation (Table 16) differed by 52.2 kJ mol^{-1}. In contrast, the computed and experimental enthalpies of formation of closely related compounds are in good agreement. An estimate of the expected enthalpy of formation can be made by use of the hypothetical reaction,

1-hexene + butane-2-one = 7-octen-2-one + ethane.

Because the functional groups are separated by two methylene units, the enthalpy of reaction would be expected to be very small; for the purpose of this exercise, let it be assumed to be zero. Given the experimental enthalpies of formation for the other compounds, the experimental enthalpy of formation of 7-octen-2-one is estimated to be −196.9 kJ mol^{-1}. From this, and from the experimental enthalpies of formation of the butenes, an estimate of the enthalpy of formation of 6-methyl-5-hepten-2-one can be made, −214 kJ mol^{-1}. This is inconsistent with the reported value of −303.7 kJ mol^{-1}. Interestingly, every difference between the calculated and experimental enthalpies of formation favors a more negative enthalpy of formation for 6-methyl-5-hepten-2-one than that expected by experiment, nevertheless the calculated result, −251.5 kJ mol^{-1}, is much more positive than that reported.

3.4.12. Fluorodinitrophenylmethane

A comparison of calculated and experimental enthalpies of formation of fluorodinitrophenylmethane and related compounds (Table 17) indicates that the reported enthalpy of formation appears to be too negative by about 46–54 kJ mol^{-1}. Evidence of this is complicated by the fact that systems with two nitro groups on one carbon atom are predicted by the semi-empirical method to be too positive by about 50 kJ mol^{-1}. Nonetheless, an examination of closely

TABLE 16. ΔH_f for 6-methyl-5-hepten-2-one and related compounds (all values in kJ mol^{-1})

Compound	Formula	Expt ΔH_f	Calc ΔH_f	Difference
6-Methyl-5-hepten-2-one	$C_8H_{14}O$	−303.7	−251.5	+52.2
1-Hexene	C_6H_{12}	−42.1	−48.5	−6.4
Butane-2-one	C_4H_8O	−238.6	−249.4	−10.8
7-Octen-2-one	$C_8H_{14}O$	−196.9[a]	−222.2	−24.7
Ethane	C_2H_6	−83.8	−75.7	+8.1
1-Butene	C_4H_8	−0.6	−5.9	−5.3
2-Butene	C_4H_8	−10.8	−21.3	−10.5
Isobutene	C_4H_8	−17.90	−25.1	−7.2

[a]Value estimated.

TABLE 17. ΔH_f for fluorodinitrophenylmethane and related compounds (all values in kJ mol^{-1})

Compound	Formula	Expt ΔH_f	Calc ΔH_f	Difference
Methane	CH_4	−74.9	−59.0	+15.9
Benzene	C_6H_6	+82.9	+87.4	+4.5
Toluene	C_7H_8	+50.0	+51.0	+1.0
Dinitromethane	$CH_2N_2O_4$	−58.9	−10.0	+48.9
Dinitrophenylmethane	$C_7H_6N_2O_4$	+35.0	+90.4	+55.4
Fluorodinitromethane	CHN_2O_4F	−235.0	−170.3	+64.7
Fluorodinitrophenylmethane	$C_7H_5N_2O_4F$	−185.0	−70.7	+114.3

related species provides circumstantial evidence of a possible fault in the experimental enthalpy of formation of fluorodinitrophenylmethane.

Using reported values, the change in enthalpy of formation by adding a phenyl group to methane to form phenylmethane (toluene) is +124.9 kJ mol^{-1}. Also using reported values, the change by adding a phenyl group to dinitromethane to form dinitrophenylmethane is +93.9 kJ mol^{-1}. Therefore the effect of the two nitro groups is to reduce the difference in energy by about 31 kJ mol^{-1} relative to that observed in going from methane to toluene. However, both changes are considerably larger than the equivalent change in converting fluorodinitromethane to fluorodinitrophenylmethane, +50.0 kJ mol^{-1}. Based on these results, the enthalpy of formation of fluorodinitrophenylmethane is estimated to be about 50 kJ mol^{-1} more positive, i.e., about −135 kJ mol^{-1}.

3.4.13. 3-Chloro-*Para*-Toluidine

The reported enthalpy of formation (Table 18) of 3-chloro-*para*-toluidine appears to be too positive by about 65 kJ mol^{-1}, as is indicated by the following reaction:

para-toluidine + chlorobenzene

⇒3-chloro-*para*-toluidine + benzene,

for which the experimental and calculated enthalpies of reaction are +62.0 and −3.0 kJ mol^{-1}, respectively. In general, replacement of a hydrogen atom by a chlorine atom at a carbon center is associated with a decrease in heat of formation. Thus when benzene is chlorinated, the enthalpy of formation drops by 28.5 kJ mol^{-1}. In contrast, when *para*-toluidine is chlorinated, the reported enthalpy of formation rises by 33.5 kJ mol^{-1}.

3.4.14. 2,4-Hexanedione

2,4-hexanedione can exist in two tautomeric forms, the dione and the enol. In the gas phase, the enol form is predicted to be about 20 kJ mol^{-1} higher than the dione. Therefore, in the following discussion, only the dione will be considered. Although the calculated enthalpy of formation of 2,4-hexanedione is only 20.5 kJ mol^{-1} higher than that observed, the value is inconsistent with those of closely related systems (Table 19). That the experimental enthalpy of formation is inconsistent can be demonstrated by comparing the change in enthalpy of formation on going from 2-propanone (acetone) to 2-butanone, and in going from 2,4-pentanedione (acetylacetone) to 2,4-hexanedione. In the case of 2-propanone, the experimental and calculated changes are 20.1 and 16.8 kJ mol^{-1}, respectively. For 2,4-pentanedione, the equivalent changes are 55.3 and 17.1 kJ mol^{-1}. Based on these results, the experimental enthalpy of formation of 2,4-hexanedione is too negative by about 35 kJ mol^{-1}.

3.4.15. 3-Phenyl Isoxazole

The reported difference in enthalpies of formation of 3 and 5 phenyl isoxazoles (21 kJ mol^{-1}) is inconsistent with the equivalent differences of the 3 and 5 methyl isoxazoles (−1.5 kJ mol^{-1}). Both calculation and experiment suggest (Table 20) that the 3-methyl derivative is less stable than the 5 derivative, and calculation indicates that the 3-phenyl derivative should be less stable than the 5-phenyl, in variance with experiment.

3.4.16. 4-Pyridinol

The experimental and calculated (Table 21) enthalpies of formation of closely related compounds indicate that the published value of enthalpy of formation for 4-pyridinol is too positive by at least 13 kJ mol^{-1} and possibly as much as 50 kJ mol^{-1}. In general, the meta position in pyridine is different from the ortho-*para* positions. This is reflected in the computed enthalpies of formation, but not in the experimental enthalpies of formation.

TABLE 18. ΔH_f for 3-chloro-*para*-toluidine and related compounds (all values in kJ mol^{-1})

Compound	Formula	Expt ΔH_f	Calc ΔH_f	Difference
Para-toluidine	C_7H_9N	+41.8	+54.0	+12.2
Chlorobenzene	C_6H_5Cl	+54.4	+53.1	−1.3
Benzene	C_6H_6	+82.9	+87.4	+4.5
3-Chloro-*para*-toluidine	C_7H_8NCl	+75.3	+16.7	−58.6

TABLE 19. ΔH_f for 2,4-hexanedione and related compounds (all values in kJ mol^{-1})

Compound	Expt ΔH_f	Calc ΔH_f	Difference
2-Propanone (acetone)	−218.5	−232.6	−14.1
2-Butanone	−238.6	−249.4	−10.8
2,4-Pentanedione (acetyl acetone)	−384.4	−402.1	−17.7
2,4-Hexanedione	−439.7	−419.2	+20.5

3.5. Unassigned Differences

Most differences in enthalpies of formation in the range ±20 kJ mol^{-1} are difficult to assign to either reference data or to calculation. In a few cases, however, there is evidence of inconsistencies in the experimental value.

3.5.1. 1-Methyl Cyclohexene

The calculated and experimental differences in the enthalpies of formation for cyclohexene and 1-methyl cyclohexene are −16.6 and +22.7 kJ mol^{-1}, respectively. The computational method is insufficiently accurate to reliably assert that there exists any anomaly with the experimental values. However, with attention now drawn to this system, closer examination of related systems (Table 22) indicates an anomaly. For *cis*-but-2-ene and for cyclopentene, the experimental change in enthalpy of formation upon forming the methyl derivative is −33.8 and −40.4 kJ mol^{-1}, respectively. This is in contrast with the change in methylating cyclohexene, −77.0 kJ mol^{-1}. Because typical errors in this type of compound are of the order of 10–20 kJ mol^{-1}, it is not possible to assign an error, if any, unambiguously. However, since most of the computed errors are negative, and that for 1-methyl cyclohexene is positive, there is an increased probability that, if an experimental error exists, it is in the enthalpy of formation of 1-methyl cyclohexene.

3.5.2. Cyclopentene Carbonitriles

One of the smallest inconsistencies detected in the experimental data concerns the two isomers, 1-cyclopentenecarbonitrile and 2-cyclopentenecarbonitrile, for which the calculated and experimental enthalpies of formation (Table 23) differ by less than 15 kJ mol^{-1}. The differences between the calculated and experimental values is more pronounced when the differences between the enthalpies of formation of the isomers is evaluated, +14.6 kJ mol^{-1} experimentally and −12.5 kJ mol^{-1} computationally. For two such similar systems, a computational error of 27.1 kJ mol^{-1} in going from the 1 to the 2 isomer would be unexpectedly large. This inconsistency can be compared to the change in enthalpies of formation in going from 1-cyclohexenecarbonitrile to 2-cyclohexenecarbonitrile, −8.2 kJ mol^{-1} experimentally and −14.5 kJ mol^{-1} computationally, and in going from (E)-2-butenenitrile to 3-butenenitrile, −17.0 kJ mol^{-1} experimentally and −16.7 kJ mol^{-1} computationally. Because of this, doubt has to be cast on the experimental enthalpies of formation of the cyclopentene carbonitriles. Similar analyses can be performed for a large number of cases where the calculated and reported enthalpies of formation differ by several kilojoules per mole. But with decreasing differences, assignment of the difference to faulty experimental or computational enthalpies of formation becomes increasing fraught with uncertainty.

TABLE 20. ΔH_f for 3-phenyl isoxazole and related compounds (all values in kJ mol^{-1})

Molecule	Expt ΔH_f	Calc ΔH_f	Difference
3-Phenyl isoxazole	+139.0	+179.5	+40.5
5-Phenyl isoxazole	+160.0	+172.8	+12.8
3-Methyl isoxazole	+35.6	+43.5	+7.9
5 Methyl isoxazole	+34.1	+38.1	+4.0

TABLE 21. ΔH_f for 4-pyridinol and related compounds (all values in kJ mol^{-1})

Compound	Expt ΔH_f	Calc ΔH_f	Difference
2-Pyridinol	−79.7	−70.3	+9.4
3-Pyridinol	−43.7	−48.5	−4.8
4-Pyridinol	−30.3	−67.8	−37.5

3.6. Tautomeric Forms

Some compounds can potentially exist in tautomeric form. In such systems, the barrier to interconversion is sufficiently small that equilibrium exists between the two forms, and the fraction present in highest concentration is always the form with the lower energy. In some instances in the NIST database[1] the higher energy form, usually the hydroxy form, is indicated. Thus, for example, the chemical structure given for 1,3,5-triazine-2,4,6(1H,3H,5H)-trione corresponds to 2,4,6-trihydroxy-1,3,5-triazine. These two structures have calculated enthalpies of formation of −564.0 and −375.1 kJ mol^{-1}, respectively, thus the hydroxy form is predicted to be 188.9 kJ mol^{-1} less stable than the trione form. The calculated enthalpy of formation of the trione agrees with the reported experimental enthalpy of formation of the trione of −564.1 kJ mol^{-1}. In cases such as this, where the difference in enthalpies of formation between the two putative tautomeric forms is large, the fraction of the higher energy form is likely to be very small, and to indicate that the chemical structure is that of the higher energy form is potentially misleading.

TABLE 22. ΔH_f for 1-methyl cyclohexene and related compounds (all values in kJ mol^{-1})

Compound	Expt ΔH_f	Calc ΔH_f	Difference
1-Methyl cyclohexene	−81.3	−58.6	+22.7
1-Methyl cyclopentene	−4.4	−23.8	−19.4
2-Methyl but-2-ene	−41.5	−56.1	−14.6
Cyclohexene	−4.3	−20.9	−16.6
Cyclopentene	+36.0	+14.2	−21.8
Cis-but-2-ene	−7.7	−18.4	−10.7

TABLE 23. ΔH_f for the cyclopentene carbonitriles and related compounds (all values in kJ mol^{-1})

Compound	Expt ΔH_f	Calc ΔH_f	Difference
1-Cyclopentenecarbonitrile	+156.4	+142.3	−14.1
2-Cyclopentenecarbonitrile	+141.8	+154.8	+13.0
1-Cyclohexenecarbonitrile	+101.5	+106.2	+4.7
2-Cyclohexenecarbonitrile	+109.7	+120.7	+11.0
(Z)-2-butenenitrile	+134.0	+143.2	+9.2
(E)-2-butenenitrile	+140.7	+142.7	+2.0
3-Butenenitrile	+157.7	+159.4	+1.7

4. Discussion

Almost all of the putative errors in experimental enthalpies of formation were identified by the large difference between the calculated and experimental values. This suggests a method whereby possible errors in the NIST data set of enthalpies of formation could be identified. When the differences are sequenced into increasing values, as in Fig. 1, the set of candidate suspect experimental data becomes localized at the ends of the curve; the compounds in these regions are listed in Tables 24 and 25. The experimental enthalpies of formation of most of the compounds in these sets are of questionable accuracy, as discussed above. Once attention is focused on one of these compounds, the apparent error becomes obvious. In some instances, insufficient evidence existed to warrant asserting that the reported values were likely to be incorrect; these are indicated in Tables 24 and 25 in normal font.

The set of examples given here is not intended to be exhaustive—a careful analysis of the 2% of compounds at the ends of the curve in Fig. 1 would likely reveal many more possible errors in the experimental reference data set—rather, the set is intended to illustrate the usefulness of semi-empirical methods for identifying possible errors in such data sets.

It is suggested that no attempt should be made to modify the values in the NIST reference data set as a result of using methods of the type described here, to do so would compromise the nature of the errors in that they would no longer be purely random. Instead, if modifying the values is considered necessary, then a new reference data set that contains both experimental and calculated values should be constructed. As an alternative, the data that are considered unreliable should be indicated as such.

The continued used of terms such as ''apparent error,'' ''putative error,'' ''questionable accuracy,'' etc., must be explained. Semi-empirical methods are parameterized using experimentally obtained reference data, therefore development of such methods depends on the existence of such reference data. Given the highly limited accuracy of semi-empirical methods, to assert that an experimental datum is inaccurate based solely on a comparison of experimental and calculated values would appear arrogant. Because of the known limitations of these methods, all arguments regarding the accuracy of experimental data were made using other experimental data, presumably of high accuracy, and not on the theoretical predictions. But these arguments assume that commonly accepted thermochemical rules are followed. The possibility must not be dismissed that some hitherto unsuspected phenomenon might be operating, and that the anomalous results reported here are in fact a consequence of that phenomenon. While this might be the case, it is difficult to imagine the phenomenon that could explain some of the experimental

TABLE 24. Largest negative differences between experimental and calculated ΔH_f (all values in kJ mol^{-1})

Compound[a]	Expt ΔH_f	Calc ΔH_f	Difference
2-Perfluorobutene	−946.0	−1565.7	−619.7
2-Tert-butyl-para-*cresol*	+207.0	−209.2	−416.2
Octafluorotoluene	−1269.4	−1492.9	−223.5
Undecylcyclohexane	−148.1	−358.6	−210.5
Pentafluoroiodobenzene	−549.0	−738.9	−189.9
2,6,6-Trimethyl-2-cyclohexen-1-one	−44.2	−232.6	−188.4
1,3,5,7-Tetraazaadamantane	+199.0	+47.7	−151.3
Difluoroacetylene	+20.9	−106.7	−127.6
Perfluorobuta-1,3-diene	−942.2	−1060.2	−118.0
Bis-(n-perfluoropropyl) ether	−3105.0	−3218.8	−113.8
1,2,3,6-Tetrahydropyridine	+140.6	+29.7	−110.9
Dodecafluorocyclohexane	−2370.4	−2470.7	−100.3

[a]Compounds with disputed enthalpies of formation are in italic.

TABLE 25. Largest positive differences between experimental and calculated ΔH_f (all values in kJ mol^{-1})

Compound[a]	Expt ΔH_f	Calc ΔH_f	Difference
Diphenyl disulfone	−481.3	−23.0	+458.3
1,3,5-Tricyanobenzene	+313.0	+509.6	+196.6
2,4,6-Trihydroxy-1,3,5-triazine	−564.0	−375.1	+188.9
3-Bromo-3-methyldiazirine	+291.0	+434.3	+143.3
Carbon monoxide	−110.5	+25.5	+136.0
Trans-difluorodiazene	+81.2	+211.3	+130.1
Diethyl malonate	−921.7	−795.4	+126.3
Di-*tert*-butyl sulfone	−546.3	−429.7	+116.6
Fluorodinitrophenylmethane	−185.0	−70.7	+114.3
5,6-Dibutyl-5,6-bis(4-*tert*-butylphenyl)decane	−351.0	−245.2	+105.8
Methyldinitramine	+53.5	+150.2	+96.7
Bis(2,2,2-trinitroethyl)-amine	+47.3	+142.3	+95.0

[a]Compounds with disputed enthalpies of formation are in italic.

results, such as the unexpected enthalpy of formation of undecylcyclohexane.

5. Conclusion

To a limited degree, the accuracy of experimental enthalpies of formation can be investigated by using semiempirical methods. Where agreement between calculated and reported enthalpies of formation exists, it is reasonable to assume that the experimental value is, indeed, accurate. Where there is a large difference, the error is likely to be either in the computational method or in the experiment; the probability of both being equally incorrect is small. If the limitations of the method are known, then the use of such methods to screen experimental data can readily be limited to those systems where the method is known to be reliable. In such cases, whenever an indication is found that an experimental reference enthalpy of formation is inaccurate, an analysis of the type presented here can be carried out to determine if further work is warranted.

Note added in proof: After acceptance of this article, several entries in the NIST database were modified, therefore some of the anomalies cited here are no longer valid.

6. Acknowledgment

This work was funded by Grant No. 1 R43 GM067327-01 from the National Institutes of Health.

7. References

[1] H. Y. Afeefy, J. F. Liebman, and S. E. Stein, in *NIST Chemistry WebBook*, edited by P. J. Linstrom and W. G. Mallard, NIST Standard Ref. No. 69 (http://webbook.nist.gov/chemistry) (National Institute of Standards and Technology, Gaithersburg, MD, 2003).
[2] M. J. S. Dewar and W. Thiel, J. Am. Chem. Soc. **99**, 4907 (1977).
[3] M. J. S. Dewar and W. Thiel, J. Am. Chem. Soc. **99**, 4899 (1977).
[4] M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, and J. J. P. Stewart, J. Am. Chem. Soc. **107**, 3902 (1985).
[5] J. J. P. Stewart, MOPAC 2002. CAChe Group, Portland (2002).
[6] J. J. P. Stewart, J. Comput. Chem. **10**, 209 (1989).
[7] J. J. P. Stewart, J. Comput. Chem. **10**, 221 (1989).
[8] J. J. P. Stewart, in *Encyclopedia of Computational Chemistry*, edited by P. von Rague Schleyer *et al.* (Wiley, Chichester, 1998), p. 2000.
[9] G. Purvis, CAChe Worksystem Pro 5.0. CAChe Group, Fujitsu America, Inc., USA (2002).